Sdibusplusfredsnotesdraftforfiling102303(asof102103)
Sdibusplusfredsnotesasfiled102302

SDI-BUS: Autonomous Corporation Auditing Tool

Frederick S. M. Herz, Warrington, PA

Walter Paul Labys, Ogden, Utah

Conversion of Provisional Application #60/420,752

References Cited

5754938	**May 19, 1998**	**Herz, et al**

Reference to Sequence listing, a table

Abstract

This document discloses an autonomous auditing tool that serves as a trusted third-party go-between for corporations, on the one hand, and shareholders, external auditors, and governmental agencies on the other. This system, SDI-BUS, uses a combination of static forms and statistical analytics to securely analyze patterns in internal managerial and financial databases for evidence of errors or wrongdoing. The results of its analysis can then be appropriately tailored and distributed to various external audiences.

Background of the Invention

In recent months, corporate scandals have become very visible signs of a disconnect between corporations' internal dealings and the information that they disclose to their own auditors or to the public in general. It is a necessary condition that not all corporate data can be revealed publicly, since the revelation of proprietary information could easily damage a company's ability to compete. On the other hand, this lack of transparency makes it much easier for errors or intentional mismanagement to occur without the knowledge of auditors or shareholders. Recent examples of such wrongdoing include
:

ADELPHIA -- Failed to properly disclose $3.1 billion in loans and guarantees to its founder's family.
CMS Energy -- Disclosed that it had overstated revenue in 2000 and 2001 by including artificial "round trip" energy trades.

Computer Associates – May have artificially inflated revenue and improperly rewarded top executives.
Dynegy -- Its "Project Alpha" transactions may have served primarily to cut taxes and artificially increase cash flow.
ENRON -- Admitted it improperly inflated earnings and hid debt through business partnerships.
Global Crossing – May have sold its telecom capacity in a way that artificially boosted its 2001 cash revenue.
Halliburton -- May have improperly recorded revenue from cost overruns on large construction jobs.
ImClone Systems -- Former CEO Samuel Waksal charged with insider trading.
Kmart -- Investigated by SEC for suspect accounting practices.
Lucent Technologies -- Adjusted fiscal 2000 revenues by $679 million, spurring SEC investigation. Agency also investigating whether vendor-financing played an improper role in its sales.
MicroStrategy -- Settled without admitting wrongdoing in SEC suit accusing it of backdating sales on tracts to meet quarterly financial estimates, among other improper revenue-recognition practices.
Network Associates -- May have hidden expenses and overstated revenue from 1998 to 2000.
PNC Financial Services -- Restated its 2001 results by $155 million after regulators raised concerns about loan transfer.
Qwest Communications – May have inflated revenue for 2000 and 2001 through capacity swaps and equipment sales.
Reliant Resources -- Admitted it inflated revenue by counting artificial "round trip" energy trades.
Tyco International – May have improperly created "cookie jar" reserves to boost profits rather than to cover merger costs; may also have "spring-loaded" earning from acquisitions by accelerating their pre-merger outlays.
WorldCom – May have used questionable methods to book sales, classify assets and account for debts it couldn't collect.
XEROX -- Fined $10 million without admitting or denying wrongdoing for inflating revenue and profits from 1997 to 2000 by including future payments on existing contracts.

Although human accountants or auditors have traditionally been used to detect such discrepancies, it is now clear that they are capable of making major errors. Whatever the cause – a lack of relevant data, corporate pressure, or an inability to "connect the dots" between complex financial dealings bordering on the illicit – there is clearly room for improvement in the performance of today's auditors.

Brief Summary of the Invention

SDI-BUS is an automated auditing system designed to overcome some of these problems. Based on the previously disclosed SDI architecture (pending patent, Secure Data Interchange, Herz et al.), SDI-BUS plays the role of a trusted third-party agent that intermediates between a company's internal affairs and interested external parties not necessarily authorized to receive proprietary information. SDI-BUS takes in a broad range of inputs that include corporate databases, news sources, and auditor's instructions; it then processes the information using a combination of machine-based and human intelligence, generating standard financial reports and scanning for patterns consistent with improper or illicit corporate activities. It then parses the information and transmits it to other interested third parties, conditioning the contents of its transmissions on the permission levels assigned to each party.

The result is a secure and trusted third-party system that replicates many of the functions of a standard auditor, but which is better able to resist corporate pressures and is capable of pulling together multiple streams of disparate information and accurately scanning them for anomalous patterns. Because the system is secure, companies can worry less about revealing sensitive information, while interested third-parties (such as government agencies or shareholders) can be confident that the resulting analyses will more accurately detect corporate wrong-doings.

Brief Description of the Drawings

FIG. 1 – Suggested SDI-BUS Structure.

FIG. 2 – Shows an example of how a Bayesian net can be used to construct the probabilistic model at the core of SDI_BUS.

Detailed Description

SDI-BUS is based on the previously disclosed SDI architecture (again, provide reference, with perhaps more description), which defines the framework for an autonomous and secure third-party data analysis system. There are, of course, a multitude of ways in which SDI-BUS can be configured; for brevity we discuss one of many potential configurations. See Figure 1 for an overview.

Inputs

Because a given implementation of SDI-BUS is responsible for monitoring a single corporation, its primary source of input data is from the corporation itself. In this example, SDI-BUS is connected securely to the set of managerial, financial, and general corporate databases that fully describe the internal workings of the target corporation, everything from the identities of employees, to sales, to the daily flows of money within the company. (What other sorts of information would we want to take directly from the company itself? – (1) corporate email and phone logs, (2) list of perks given to executives, (3) business trip reservations data (e.g., a series of flights to the Cayman Islands might indicate offshore activity), (4) document management system.) Note that a secure transfer protocol is utilized to minimize the risk of information leaks.

SDI-BUS also accesses a wide range of external information; one example would be Reuters' international business news wire; another example would be the public databases that describe corporate activities (e.g. mergers, identities and biographies of board members, insider trading statistics, subsidiaries, offshore connections, ownership networks, etc.). (what other sorts of external data? (1) data from banks involved with handling the company's money, (2) industry specific data – for example, when dealing with energy companies it would be good to have ties to energy spot market feeds in order to detect suspicious trading patterns, (3) ties to proprietary SEC/DOJ databases revealing previous infractions of companies and individual company officers.) **(4) personal and professional dociers assembled from multiple public and private databases. C0-pending patent application entitled, "System and Method for Pre-screening Potentially Litigious Patients", Herz, et al., incorporates a plethora of different types of data collected from almost any type of public and/or private database containing records relating to a particular subject individual of interest. Statistical analytical methods, in turn, are able to predict from such data patterns in the data which are consistent with a variety of types of behavior which is in some way of concern. While the system as described primarily focuses on the future or predictive likelihood of a certain negative behavior to be performed (e.g., litigation, fraud, etc., when implemented as part of the presently disclosed system and methods (which incorporates any/all of the presently disclosed input variables), the same analytical methods may predictively determine the likelihood that a particular undesirable behavior had already occurred.**

Analysis

Once these various strands of data are gathered within SDI-BUS, a variety of analytical methods are applied: these range from static scripts or templates, active statistical analysis (there are many possible approaches already given by the literature), and occasional human intervention. See Section 3 for more details.

The products of this stage may include the following:
* Standard Business Audit Reports (i.e., very similar to what a traditional auditor would produce for a company): basic breakdown of money stocks and flows within a company.
* Business Strategy Reports (similar to what a strategy consultant might produce): given full access to a

company's data it is very likely that inefficiencies or missed opportunities in the company's operations will be detected by SDI-BUS. It can produce a report recommending actions that the company can take to correct these problems and increase profits.
* Fraud detection: This report details areas within the company's operations that have exceeded a certain probabilistic threshold of suspicion and have been flagged for follow-up by an expert human auditor.

Feedback -- Adjustments

In the preferred implementation of the system, an independent human auditor is allowed to monitor the reports being generated by SDI-BUS. In addition to having the ability to oversee the generation of reports, this auditor will also have the ability to tune various settings within the system. These alterations include:

* The auditor can alter the probability thresholds used in fraud detection (a company that sells scuba gear might have a very valid reason to do business in the Cayman Islands – an auditor might therefore greatly increase the threshold needed to trigger the "hiding profits offshore" flag).

* The auditor can write new templates specific to the target company and tailored to the auditor's particular focus. These templates can then be uploaded to SDI-BUS, which will include the new templates in its ongoing analysis. An additional layer of protection will prohibit SDI-BUS from returning any proprietary information to the auditor, although it may be the case that non-specific data measures, such as randomized aggregates, are allowed.

* The auditor can upload modules incorporating entirely new statistical or analytical methods. Depending on the complexity and on the variables requested, these modules may need to be security checked before they are allowed into the SDI-BUS system.

Dissemination of Results

When reports and analyses are complete, SDI-BUS disseminates its results on a need-to-know basis to a collection of interested third parties. Thus, while an internal company auditor might be given a report that includes proprietary information regarding certain sales figures in the Cayman Islands, outside shareholders wouldn't be given access to this inside information, and might rather be told simply that the company's current fraud-detection index is "high".

Interested parties receiving final reports include the following:

* Company Auditors
* Government Agencies: Securities and Exchange Commission, Dept. of Justice, etc.
* Investors/Shareholders

3. Analytical Methods

SDI-BUS is capable of using a wide variety of analytical methods, which may be used in different combinations or settings depending on the target corporation and on the analytical objectives. These methods include, but are not limited to, the following:

Templates/Rule Base

The simplest type of analysis is done by templates, or fixed rules, that are stored internally in SDI-BUS's rule base. These generally embody basic rules of accounting and corporate governance, and can be quickly fitted to data as it is loaded in.

For example, one could set up a rule to track executives previously suspected of setting up offshore shells:

xi = binary variable flagging previous conviction by DOJ for executive i. (0=no conviction, 1=conviction).
yi = number of trips to Caribbean in last year.
zi = number of phone calls to Caribbean in last year.

IF (xi) AND { $yi > 3$ and $zi > 5$ } THEN OFFSHORE_FRAUD_FLAGi=1 ELSE OFFSHORE_FRAUD_FLAGi=0.

In operation, SDI-BUS will run a battery of such simple rules against the target data.

Statistical Analysis

There are, of course, many more sophisticated methods of analysis, such as the various well-understood and publicly documented statistical techniques popularly used for data mining. In the preferred implementation, one of the analytical approaches used by SDI-BUS will involve Bayesian networks, which allow a complex web of inputs to successively influence the probability distribution of a final outcome. In this case, it uses the inputs to calculate a final probability distribution for the following event: is fraud occurring?

The simplest types of Bayesian nets are directed acyclic graphs that encode conditional probabilistic relationships between various event nodes. Realistically, it is likely that the complex graphs used for this application will include cyclical elements (in other words, more than one semipath may exist between any two nodes); this makes the calculation of conditional probabilities more complex, but still within the realm of solvability (using known state-of-the-art statistical methods).

The network can be constructed by human domain experts who understand the many factors involved in business and accounting fraud, as well as their causal linkages. These factors include those things that would impact the probability of fraud occurring (for example, recruitment of chief executives with histories of unsound financial dealings), as well as those things whose probabilities of being observed are impacted in turn by the occurrence of a fraud (for example, a sudden drop in stock price or sudden increase in insider selling).

Because the central event of interest – the occurrence of fraud – may not be directly observable in its early stages, our calculation of its probability will be heavily conditioned on those factors which are directly observable.

Once the network connections are established, the conditional probabilities for the event nodes must be defined. Although it is likely that most of these will again be constructed by human experts, there are well-known machine learning methods that would allow the probabilities to be calculated directly from a training data set. Certainly, once the system has been in operation for some time and enough data has been collected, the overall accuracy of the network could be improved by training it on the new data.

Figure 2 shows an example of how a Bayesian net can be used to construct the probabilistic model at the core of SDI-BUS. Note that this is only one embodiment, and that different structures and/or different variables may be built into it. For this example, we indicate observable events as boxes and unobservable events as ovals. Note that the direction of the arrows indicates which events have an impact on the probabilities of the events that follow. The oval representing the fraud event lies in the middle of this network: some events feed into it, and it feeds into other events. The information contained in both the "parents" and the "children" of the event will be used in the calculation of its conditional probability.

At the top of Figure 2 is a row of observable events that may occur before a fraud actually takes place – the target company may begin to employ individuals knowledgeable about methods for committing fraud (it is likely that even if such individuals do not have criminal records, they will be tied through newspaper accounts and press releases to unsavory past fraud events, companies associated with fraud, or other individuals who ended up facing charges for fraud). The company might also start opening offshore branches or accounts. Finally, there may be preliminary announcements of concern by the Department of

Justice or the SEC regarding the company's activities.

If and when fraud occurs, it will probably not be observable from the outside. However, the consequences of the fraud may be observable: for example, the company's announced financial results may be spectacularly better than those of other companies in the same industry.

Although fraud may not be observable to the outside world, it will definitely be observable to company insiders; if they are not happy with the event, they may panic and take defensive actions. Although the insider panic will not be directly observable to the outside world, it will likely cause certain observable events. These might include an increase in insider stock sales, an unexpected drop in share price, incidents of document shredding, and a wave of resignations by ethical executives who wish to distance themselves from the company.

Once the weights within the Bayesian network are configured (possibly by human experts using past events as guides), the network itself is ready to be used as a fraud monitor. At regular intervals (depending on the company, anything from daily to monthly) values for observable events are calculated and fed into the network, which then computes a distribution for the probabilistic event that a fraud has occurred.

Human Expert

Finally, there is a provision for human experts to be included in the loop (with proper security provisions) when especially sensitive or complex issues arrive. Most often they will be called in to verify or dismiss potential fraud when SDI-BUS's automated analysis flags an irregularity.

[redacted]

Scamster Fraud Types

1.Pump and Dump schemes – Bogus or over-hyped company stock is over-hyped over the Internet by the scamster so stock value becomes artificially inflated. He then dumps his stock to these new over-zealous investors.
2.Claiming to have invested money into non-profits but actually investing it into foreign companies for profit in order to both conceal profits and avoid taxes for these profits without detection.
3.Within a company, employees making requisitions of purchase orders and authorizing delivery of equipment with the intention of stealing it.
4.Particularly in financial services industry, providing business or personal loans to executives, colleagues, employees or families thereof (including "political favors") without assurances or even intent to act in good faith in repaying the loan (e.g., allowing the company to go bankrupt). SOA prohibits employees or executives from receiving such loans.
5.Compellng (illicitly incenting) a competitor's manager not to drive the other companies out of business (which is its job to do so, e.g., a new grocery, convenience or pharmacy chain store whose primary mission is to drive out of business the smaller longer established private competitors).
6.Insider trading – under SOA insider trading schemes (e.g., ImClone, Waksal and Martha Stewart) would have to give detailed disclosure of timing and quantity of stocks sold as well as significant sales by other people (e.g., Waksal's family and friends), and this data inherently would have to be disclosed as linked to particular

harmful events to the stock value (i. e., failure to receive SOA approval of the company's primary drug for cancer).
7.Moving funds into international divisions for purposes of performing fraud in these less regulated markets – SOA requires the same regulations for foreign divisions of American companies as those regulating its domestic divisions.
8.As in the Enron scandal, over inflating earnings, underestimating debts, such that company performance appears better than it is and so executives can cash in on additional stock options (and presumably then cash out of their equity before the true status of the company affects the actual stock price). This is one example of an executive's opposing interest to achieve short-term gains, or perceived short term gains at the expense of the company's long term valuation (this is more of an incentivization optimization problem regarding stock option incentive formulas in response to performance criteria).

SOA strongly places additional responsibility and burden on upper level management even for fraud and misdeeds at the lower levels. It also is designed to incentivize whistleblowers in order to form a more distributed economic structure for self-policing.

Auditing Concepts
1.Fraud is typically associated with the movement of funds from legitimate to illegitimate uses in actuality and on reports. Clever fraudsters are able to fuge numbers such that the uses and destinations of such funds appear legitimate. In certain cases the actual numbers may be greater than or less than the reported numbers for income receivables, payables, expenditures, etc., and the nature of the allocations/uses and sources of these funds. If there is sophisticated tampering of these records by a clever fraudster, the continual use of electronic work flow applications in which templates can be automatically or manually filled could be used to statistically flag anomalous behavior such as the above. The actual internal fund transfers between the organization and/or financial institution could conceivably provide the process for which these automated templates could be filled thus enabling better tracking of monetary inflow/outflow and money handling within and without the organization. A certain degree of the auditing process is based upon providing explanations of financial data including anything from the reasons for certain financial activities, expenses, receivables, loans, forecasts over actuals, missing data records, etc. It may be possible to custom construct certain templates in which natural language is parsed, the templates are automatically filled and natural language processing techniques are used in order to identify certain particular anomalous patterns by which a notification alert may be triggered and, for example, additional human and/or automatically generated questions could additionally be presented to the party(s). A decision tree or other similar hierarchical querying scheme could be used to immediately query all relevant individuals associated with any internal conditions or events which are even slightly anomalous , inconsistent and/or potentially suggestive of hiding or concealing of certain financial information or events which may correlate with that information. An associative web of data, people and events may also be collected and aggregated

over time in order for link analysis to occur, for example, tracking of an event to certain individuals and/or web of associates or family members associated with these individuals. This can be useful both for purposes of collecting factual evidence from multiple sources about and even on individuals or for performing statistical analysis from the plethora of personal available data as to the statistical probability the likelihood of such individual(s) to actively engage or associate themselves with suspicious activities.

The system would be more effective in identifying the probable likelihood of certain suspicious activities or events by monitoring a more complete collection of the sources and channels by which an individual in a company who is motivated to commit fraud would necessarily have to utilize. For example, SDI Scam could opportunally monitor the business activities, transactions and investments of a given individual within a company (and ideally those of his/her family and/or friends/associates), particularly if certain statistical derived suspicions have already been raised.

Describe automated regulatory compliance and autonomously provided attestation using rules based on automatic classification of behavior, some template based language inputs and a rule base characterizing the various regulatory requirements.

Describe how to limit a plethora of SOA compliance criteria without shareholders being revealed the exact nature of each of the various regulations complied with.

8

In operation, SDI-BUS will run a battery of such simple rules against the target data.

Overall, rules will fall into three general categories:

1) Rules generated by laws, regulations, or conventions.

Laws, regulations, and conventions tend to be fairly fixed, and it is straightforward to convert them directly into fixed rules or templates. For example, generally accepted rules of accounting often specify particular relationships between financial variables. If a particular business is required by its industry to maintain a debt no larger than 33% of revenue, such a requirement can be directly translating into a rule based on financial variables:

Rule A: debt <= 0.33 * revenue

If this inequality is violated, the rule is "triggered."

2) Rules created by human experts

In other cases, rules may be based on the intuition or experience of human experts, who can write rules directly in a form of pseudocode. Such rules may still involve strict relationships between known variables, but the nature of the relationships and the identity of the variables may be very diverse, limited only by the imagination of the creator. For example, a Human Resources expert may believe that, to avoid conflicts, messy and neat employees should never be assigned to the same workspace. Given an index of "messiness" (perhaps derived by scanning work directories and estimating level of organization; it may be the case that messy employees don't mind having large numbers of heterogeneous or outdated files scattered across a few directories), a rule for employees A and B might be encoded as follows:

Rule B: abs(neatness_A – neatness_B) <= difference_threshold

Suppose a company is in the process of assigning office mates, based on various factors. If such a rule is triggered by HR software the director may opt not to assign employees A and B to the same space.

3) Rules estimated by statistical analysis

Finally, there are known statistical techniques for deriving rules (often related to classification) from sets of data. Although such rules may not make intuitive sense to human observers, they often embody complex and significant patterns in the data. For example, an insurance company may run an analysis of their customers and determine that combinations of car color and cell phone ownership may correlate with bad driving habits. A rule may appear as follows:

Rule C: 0.223*car_color_is_red + 0.998*cell_phone_minutes_per_month <= 23.5

If such a rule is triggered (that is, if the inequality is violated) the company may have reason to believe that a given driver is an insurance risk.

Combination of Rules

The combination of rules that is used for a particular installation of SDI-BUS will depend entirely on the goals of that particular enterprise. A business that requires strong compliance with industry standards may select a set of rules that simply reflect coded regulations, and these rules might only require access to the company's accounting data. On the other hand, a business worried about internal fraud may select a wide range of statistically-derived rules that are applied to all available data, including employees' emails and personal files.

Once a set of rules is chosen, each rule is assigned a particular weight to reflect its relative importance. When the full set of rules is applied to target data, individual rules may or may not be triggered in response to the observed data; the sum of the weights of the rules that do get triggered represents the overall response of the rulebase.

The significance of the rulebase response depends entirely on the threshold of the user. In cases where all rules are significant, any nonzero response may require user attention – in short, the reaction threshold is zero. On the other hand, if the requirements of the installation are such that the triggering of a few minor rules is inconsequential, the user's threshold will be some larger positive number.

An important aspect of the rulebase is that it is fully dynamic: rules may be plugged in or plugged out as necessary. Moreover, rules themselves may be modified to reflect changes in regulations, for example, or when statistically-derived rules are re-estimated on updated data.

Most important, the user always has access to the rules themselves as well as their weights. Hence, the rulebase may be dynamically updated as the goals or thresholds of an organization change over time.

Within such a set, each rule is assigned a particular weight. The ensemble of rules is then applied to current data, and the weights of those rules that are triggered are combined arithmetically.